UNITED STATES DISTRICT COURT

FOR THE WESTERN DISTRICT OF NORTH CAROLINA

CHARLOTTE DIVISION

Case No. 3:22-cv-0191

KANAUTICA ZAYRE-BROWN, )
)
Plaintiff, )
)
v. )
)
THE NORTH CAROLINA )
DEPARTMENT OF PUBLIC )
SAFETY, et al., )
)
Defendants. )

Deposition of FAN LI, Ph.D.

(Taken by the Plaintiff)

Raleigh, North Carolina

Friday, August 11, 2023

Reported by: Marisa Munoz-Vourakis -
RMR, CRR and Notary Public

Okay. So what this study is proposed to do is try to use real world data, real world data means like observational study or the historical data, historical clinical trial or current like electronic health record, use that data to combine that data, this clinical trial data, and that is, yeah, so there are a lot of challenges, methodology challenges there, and then that's what this grant is supposed to do to develop those methods.

Q. Okay. So it's about finding methods to be able to use that real world data?

A. Well, combination, basically clinical trials. Like you can have clinical trials. You can also have external data. How do you properly combine them?

Q. Okay. And can you tell me about the second grant that is listed?

A. Oh, the second one, that's a funny one. So the second one is this Covid-19 Enhancement. So this actually is better to read it. So second and fourth actually kind of connected.

So the fourth one is my, again, I'm just -- I'm not the primary investigator there. So if you look at that it's called Co-I, it's a coinvestigator.

So the number four, that was the kind of

the mother grant with my collaborator Laine Thomas got the grant from PCORI. PCORI stand for Patient Centered Outcome Reference Institute.

So that was the primary grant, and then when Covid-19 hit, the agency had more money, and so then we asked, like we basically ask for supplement to do additional methodology work. So that's why you have this number two. So that's actually just, again, a supplement of the grant number four.

Q. Okay. And at a very high level, what was grant number four?

A. Oh, again --

Q. What was the goal?

A. The goal is try to develop statistical methods to improve the design and the conduct of the observational study, particularly this type of subgroup analysis. So high level is the observational study, so the methodology and the design and the methodology. There are challenges, and we propose new methods to solve those challenges.

Q. Okay. So going back to the substance of your report, can you summarize your conclusion in your report for me?

A. Absolutely. Give me a minute.

So I will just read it.

Q. Actually, I don't want you to read it. Can you summarize it for me?

A. Yeah, I know. So I think I should -- okay, I should start with what my assignment was. So I was asked to assess whether the assertions made by Dr. Ettner and WPATH, the specific assertions made by them are supported by the reference they cited. And my conclusion is I do not believe that their assertion, specific assertions point by point are supported by those -- well, supported by the body of literature they cited.

In other words, I will say that the evidence cited by Dr. Ettner and WPATH do not provide reasonable or sufficient support for the assertions, specific assertions.

Q. Okay. And what do you mean by reasonable?

A. What do I mean by reasonable is reasonable judged by me, my expertise as an expert in the so-called comparative effectiveness research.

Like I'm an expert, so as a statistician, my expert is actually called inference. Well, the inference is essentially to determine what kind of evidence needed to establish the effectiveness, effectiveness or safety of a treatment, and that's my expertise. And based on my expertise, after I review,

the references cited by them and look at the specific assertion, I don't think that they provide support, provide -- it's just not enough.

Like if I'm a medical -- if I'm a reviewer of a medical journal, if you submit a paper with those assertions and those references, I would reject it on the ground that the evidence is not enough.

Q. Okay. And so how much is enough?

MR. RODRIGUEZ: Objection, vague. You can answer.

A. Yeah, I think your question is very vague, and it's beyond more than being vague, it's also this is all very case by case.

So I read a paper, for example, if it says if you want to establish the effectiveness of say a vaccine, and then I would expect to see evidence, clinical evidence from a randomized trial or well-designed observational study, and if it does not provide that, or the study does not have good quality, for example, those retrospective study assay, then I will judge it's not enough. So it's highly case by case.

So, again, that's why I was very specific. So in my report, I provide not a blanket statement. I went case -- point-by-point the assertion and exam

one-by-one and say why I believe that is not enough.

So in order to answer that question, I think it's better for you to actually go through the details.

Q. So at the bottom of page four in your report, so at the end of the second line, you said: I conclude, to a reasonable degree of statistical certainty, that these studies failed to provide rigorous and consistent statistical evidence on the benefits and quality of life and well-being of sex reassignment surgery. Do you see that?

A. Yes.

Q. What do you mean by statistical certainty there?

A. Oh, statistical certainty, again, because I'm a statistician, right, I look at the papers, those papers and the papers, that actually a big quantity of papers, 80-some studies, and the statistical certainty means that my statement is based on the kind of data, the data of those 80-some papers I reviewed.

So that's what I meant by statistical certainty. It means that I -- my judgment or my opinion is based on data, again, data in this case, those papers.

Q. And how did you calculate that statistical

certainty?

A. I don't calculate that statistical certainty. That is a phrase, and, yeah, you cannot calculate. But I based on my expertise as a statistician, my expertise in this field.

Q. Okay. So the phrase is not that there is certainty supported by numbers, but rather that you are confident and you relied on statistics?

A. Yes.

Q. Okay. And what do you mean by rigorous and consistent statistical evidence?

A. Oh, rigorous and consistent, okay, so rigorous is, again, rigorous is based on the strengths of the study. So my judgment of the strengths of the design quality, the quality of the study, that means rigorous.

So in science, we always say whether this study is conducted with vigor. So things with more vagueness, or like with a lot of authenticity and noise, we call that that's not rigorous.

And consistent, that is consistent means if you read through my report, you will see that I review this 80-some documents and I divided the papers into different categories by design. So they are -- one of my studies, they are prospective studies, and they are

retrospective studies, and the prospective studies, I cited many of them, there are not too many, there are five of them. They provide mixed results, mixed results about the benefit of the sex reassignment surgery on quality of life and well-being of the -- well, quality of life.

So they are mixed results.

And so, yeah, by consistent, I mean all the studies suggest, point to the same thing, and I don't see it here.

So that's what I meant by rigorous and consistent, and the documents, the papers that I reviewed do not provide rigorous or consistent evidence.

Q. Okay. And only randomized control trials qualified as rigorous?

A. No, I didn't say that. I never say that in my opinion. I mentioned in my opinion randomized study is to establish the treatment -- effectiveness of a treatment. They are hierarchy of study designs. The gold standard is randomized study, but I didn't say that's the only one or that it must. I said is the gold standard is randomized study, and then the second one, when that is not available, you can resort to observational study, but it need to be well designed.

The best quality of observational study design is prospective study, this before and after comparison, and that is doable, that is visible, and that is reported in some of this 5 of these 80-some studies.

And the last, the worst quality of all of this is the retrospective study results, before/after comparison, and that was the bulk of the study that was cited.

So that was my argument. I never said RCT, randomized trial is the only -- is the only one that is rigorous.

Q. Okay. And you mentioned kind of the mixed results. How many or what percentage of studies have to show the same results for them to be considered consistent?

A. You cannot make that judgment. That's not a good question, because you can have a bunch of very low quality ones, one hundred -- and all of them are very deeply flawed methodology wise, and one hundred percent of them point to -- point to one direction.

And then if you have a small bunch of high quality study, they all point to the other one, other direction. In that case, I would take that small bunch of quality over the large bunch of low quality studies.

So you can never say -- you can never say that oh, how much, what the percentage has to be, what's the percentage of the signals has to be in the studies, then it can be regarded as enough. You don't do that. You judge by the first, so their order in the judgment, so their order in the criterion. Well, the first order of business or the criterion would be the quality of the design.

So, yes, if there are two randomized study, if there are only one of two randomized study, show me one direction. And then I have one hundred very methodologically flawed low quality retrospect study shows the treatment from the other direction, I probably will trust -- well, not probably, I will trust, and FDA will trust a randomized study, take the randomized study result, which is much more -- consider that is much more reliable or valid evidence.

Q. Okay. So stepping away from the quality of the design, we're talking about consistency. You've concluded here that they are not consistent?

A. Yes.

Q. So I guess I'm trying to understand where that line is? What makes it consistent versus what makes it not consistent? And how do you know here to conclude that it is not consistent?

A. Yes. If you go to my expert report, consistent -- it's hard to find this, let me see. I think for the assertion -- let me see, the WPATH assertion one. So it's on page 12.

Q. Okay.

A. So I state very carefully, so first of all, there's no RCT, but that is not end of the world, because they are -- actually they do have five prospective studies. And so I wrote this very clear. I don't need to read it, but I cited to see that, you know, they have mixed results for them to Lindqvist -- if you read page 12, the second to the last paragraph: Lindqvist, et al, found mixed results on the effects of SRS on quality of life. Specifically, they found that comparing to before treatment, quality of life is better one year after operation but worse three and five years after operation.

So that is -- so that's one. And also if you flip to the next --

Q. So is that finding sufficient for you to find that there's not consistent statistical evidence?

MR. RODRIGUEZ: Li, she's still answering your question.

A. Yeah, I'm not done yet.

Q. Okay.

A. Because that is several -- so there are actually five prospective studies cited here, and those are of higher quality than the retrospect study.

And then here, one-by-one I actually specific tell what their findings. And not -- the Lindqvist one I read. I don't know whether I should read all of it, but you can read, you can read it yourself to say that the other one, so let me see, like the --

Q. I don't need you to read the actual citing. That's not my question. So maybe I can rephrase. Would that be helpful?

A. Yeah, again, before you rephrase, I just answer that you say not consistent, yes.

So what I wrote here is the five higher quality of higher quality observational study, they find mixed results. They don't find all that it's helpful for quality of life, and this is every paper is like, and, yes, that's my answer.

Q. Okay. So stepping away from this report, I guess I'm trying to understand how you determine when it's consistent or when it's not.

And so is it that, you know, one study out of five, is it -- you know, I don't need an exact percentage, but is the expectation that all of the

studies are consistent?

MR. RODRIGUEZ: I'm going to object to form. You can answer.

A. Oh, again, this is same kind of problems I have with your previous question is vague. It's -- well, let me say this: So consistent, again, so I have -- so, again, you have -- so this is already you cannot just take the consistent out of context.

So here there's five prospective studies, right. And I think at least three of them, the results are saying that the result is not beneficial or some of the result are beneficial, some of the outcomes are not beneficial. And so that is just -- that's inconsistent. I mean, there because you have five different -- you have five studies, and the five studies say -- have very different conclusions. And that is inconsistency. I cannot name whether there's one out of five or three out of five because that is inconsistent. Like here, at least three out of five they are showing mixed result. And also this is just mixed result, and those result are different from the direction that from the large body of low quality retrospective studies.

So that's what I meant by not consistent, because you have both retrospective study and

prospective study and observational studies, and they are of different results.

Q. Okay. Going to page 25 of your report, to the conclusion.

A. Yes.

Q. At the end, you state your opinion to a reasonable degree of statistical certainty that the studies cited by Ettner and/or WPATH and reviewed in this report simply do not provide reasonable support for the assertions made, and then it continues.

A. Yes.

Q. What do you mean by reasonable support?

MR. RODRIGUEZ: Asked and answered. You can answer.

A. I think I answered your question before.

A reasonable -- I already said a formal reviewer of a medical journal, and you submit things, you make assertions of this, and then you submit your -- and your paper or your report, you submit evidences. And based on like those -- the reference they cited, I would reject it, because I don't -- I already explain many times why I think that's -- the quality is low. The better quality of study that is available show mixed results. The lower quality ones, they are subject to a lot of methodology flaws. And

even their own expert, like even their own -- many of the papers they cited, I reviewed, call for better design studies. In that instance, I would not think -- I would not regard the evidence that they showed or in this case it's the reference they cited, provide reasonable or provide support, sufficient support for their argument.

Q. Okay. And this -- so the standard of reasonable support that you're using is sufficiently supported for you to accept it for publication in a medical journal?

MR. RODRIGUEZ: Objection, mischaracterization of testimony.

You can answer.

A. No, I use that as analogy, but I'm just saying as a scientist, how do you establish evidence for something? Okay. You actually go to the, like in this case, you want to establish the effectiveness of medical intervention. And there are ways of study -- there are different ways, study designs you can do. You can conduct those studies. And then you can analyze the data.

And, again, as an expert, I am involved in many of like this type of study, not about gender dysphoria, but just to kind of establish the

effectiveness of treatment. And my expertise, my experience tell me, based on that, I judge that what they cited here do not support their assertions. So it's overstretched.

Q. Okay. Can prospective observational studies provide reasonable support?

A. So the answer to that is not black and white.

So the best one, if possible, there would be a randomized study that's gold standard, but that is -- when that is not available, if you --

Q. Okay. Continue.

A. So that's the gold standard. I'm just stating a consensus in the field. When that is not available, then you have, yes, then you try to do the best quality observational study.

So, yeah, prospective study can be used, prospective study can be used as part of the evidence, but whether that can be viewed towards -- can be taken as the foundation for treatment recommendation, that's not a question that I can answer, because I'm not a medical doctor. But I can just tell you that the prospective study, yes, it is well conducted. If it's repeated many times with consistent result, then, yes, I believe that that would -- some medical doctors can

take that as a strong evidence. But whether that can -- again, who made that decision, I don't know, and I'm not going to -- like that's not the question I'm going to answer, like I'm supposed to answer, because that's not my expertise.

Q. Okay. Just to clarify on that, you said you're not opining on whether treatment recommendations can be made from this?

A. Well, yeah, so treatment recommendations, so evidence needed to make treatment recommendation is a different point, it's a different question from how do you establish the effectiveness of a treatment? Those two are different things.

So my assignment as a statistician, I can answer the first. I can answer the later question. I can answer the question of how do you establish the effectiveness of a treatment, but that is not equivalent to how do you make a treatment recommendation, because making a treatment recommendation, from my understanding, involves a lot of, you know, other considerations.

So -- but my, again, my scope is on whether the treatment, it's not treatment effectiveness. Arguably, that's one of the most important point consideration in making treatment recommendation.

Arguably that's the most important one, but that might not be the only one.

But, again, my scope is to answer the question, the more focused, specific question about whether you can establish the treatment -- the effectiveness of a treatment.

Q. Okay. So what is the significance of your conclusion that the studies failed to provide reasonable support for the assertions made by Dr. Ettner and WPATH?

MR. RODRIGUEZ: Objection, vague and form. You can answer.

A. Yeah, what do you mean by what is the significance? The significance in what aspect? In terms of what?

Q. So, okay. So you're talking about -- your question is about establishing the effectiveness of the treatment. So is your conclusion that the effectiveness of the treatment has not been established to a degree of statistical certainty?

MR. RODRIGUEZ: Objection, mischaracterization of the report, and her previous testimony as to what her conclusions are. You can answer.

A. So, again, I don't quite understand what

you are trying to ask here.

So my conclusion is, yes, based on my expertise and the document I reviewed, I conclude that the body of literature they cited do not support their assertions.

And you asked me is that significant? Again, sorry, I don't understand. What do you mean? What do you want me to say? I mean, what do you mean? Can you phrase this again?

Q. Yeah. Well, let me ask a follow-up question on that, because you're saying that the evidence does not support the assertion. And in layman's terms, I'm wondering if support has a different meaning, right, because is it that they don't support it with vigor, or that they don't support it at all?

A. Oh, okay. Oh, it's just they make assertion -- if you look at the assertions, some are very -- okay, let's go to specifics.

So, for example, some of the assertions I -- yeah, what do you mean by support? You want me to elaborate on that. I can do that.

So, for example, page 11 on my report, WPATH assertion one: There is strong evidence demonstrate the benefit of the quality of life. And so

I focused on this strong evidence.

So they claim this is strong evidence. I don't think that has strong evidence, okay, so that's what I meant by no support.

And the second, to give you another example, I should probably go to Dr. Ettner's assertion, let me see.

Q. And we will go through these one-by-one later, so you know.

A. Okay. That's even more fun. But, yeah, that's what I meant. Like I was very specific in this specific point-to-point discussion.

So when I say do not support, I meant, yeah, point-by-point, yeah, it's better point-by-point, then you go and see the problems. I mean, they do not support. It's often exaggeration or overstatement or sometimes just factually mistake, factual mistake. Sorry, go ahead.

Q. Are you saying that the WPATH standards of care are wrong?

MR. RODRIGUEZ: Objection, mischaracterization of the report and the testimony. You can answer.

A. No. It's a long document with many different opinions. I cannot just say it's wrong, no

it's not. I was talking about the assertions they made are not supported by -- well, again, as I just described not -- yeah, supported by their -- well, in other words, that it's often an overstretch, exaggeration, over characterization of the -- of things, of opinions. I didn't say it's strong.

Q. Are you saying that the WPATH standards of care should not be followed?

MR. RODRIGUEZ: Objection, mischaracterizes testimony. You can answer.

A. No, I didn't say that. I don't have the expertise to say things like that.

Again, my opinion is very focused on specific -- the assertions to counter or to exam the specific assertions, and I didn't make any statement about whether they should follow WPATH.

Q. Are you saying that WPATH should not rely on the studies that are cited in your report?

MR. RODRIGUEZ: Same objection, mischaracterization of the evidence and the report. You can answer.

A. I didn't say. Again, WPATH, what kind of document they want to decide on is they decide. It's not my opinion about. My opinion is about they cite, they decide, they cite a bunch of papers, references.

And, again, I try to judge that, whether the reference they cited support their assertions. So that's a different concept.

Q. Okay. Dr. Ettner stated in her report that the standards of care for treatment of gender dysphoria are currently set forth in the WPATH standards of care.

Do you disagree with that statement?

MR. RODRIGUEZ: Objection, outside of the scope of this witness' opinion. You can answer.

A. Yeah, yes, I agree with Orlando that is outside my scope.

Again, my job is to exam the specific assertions of Dr. Ettner and WPATH, and what you just asked is not part of the assertions I was asked to provide opinions on.

Q. Okay. So you are not providing an opinion on whether the standards of care for the treatment of gender dysphoria, or you are not providing an expert opinion on the standards of care for treatment of gender dysphoria?

MR. RODRIGUEZ: You can answer.

A. No, I was not -- I'm not providing opinions on that.

Q. Okay. And Dr. Ettner also stated that the

WPATH standards of care are the internationally recognized guidelines for the treatment of persons with gender dysphoria and informed medical treatment throughout the world.

Do you disagree with that statement?

MR. RODRIGUEZ: Objection, outside of the scope of the witness' opinions.

You can answer.

A. Yeah, that's outside -- that's not what I'm asked to write opinion on. It's outside the scope of my expert opinion.

Q. Okay. Dr. Ettner also stated that the American Medical Association, the Endocrine Society, the American Psychological Association, the American Psychiatric Association and a host of other entities all endorse treatment protocols in accordance with the standards of care.

Do you disagree with that statement?

MR. RODRIGUEZ: Same objection, outside the scope of the witness' opinions.

You can answer.

A. Same answer. It's outside my opinion, my report.

Q. Do you think that all of those medical associations should not endorse treatment protocols in

going to be considered low quality evidence?

MR. RODRIGUEZ: Objection, lacks foundation. You can answer.

A. As I said I don't -- I'm not familiar with the GRADE system. So I don't know whether that is what you just asked, whether observational study always considered low or high. I repeatedly said that my expertise, my experience with observational study, even with -- observational study is a very vast, a range of studies. Some designs are of higher quality. Some are lower quality. But, again, where do they fall into the GRADE system, now who decide that? I don't know.

Q. Give me one minute.

(Pause.)

Q. Okay. Do you always consider observational studies as low quality evidence?

A. No. I said clearly in the world of comparative effectiveness research, there's a hierarchy, RCT randomized study is the best. Observational study, there are good ones. There are bad ones. There are high quality ones. There are low quality ones. I don't blankedly(sic) say observational study are all of low quality. I never said that.

Q. Okay. Does a study being low quality, in your opinion, mean that it does not have value?

MR. RODRIGUEZ: Objection, vague and ambiguous. You can answer.

A. That's not my opinion is about. I don't stretch anything. I just purely say that when I say low quality, I meant there are flaws. There are serious flaws in the study design, that has rendered the conclusion be unreliable or subject to noise. And so more studies are needed.

So I didn't say that whether they have value or not. That's not up to me to judge.

Q. Do you have an opinion on whether low quality studies should be used in treatment recommendations?

MR. RODRIGUEZ: Objection, outside the scope of the witness' testimony. You can answer.

A. I don't have opinion on that.

Q. Okay. So going back to your report --

A. My report?

Q. Yes.

A. Okay.

Q. On page 9.

A. Yes.

Q. You discuss prospective studies and retrospective studies. Those are both types of

observational studies, correct?

A. Oh, caveat, yes. So observational study can -- so randomized experiment also belong to the broad spectrum of prospective study. But -- well, because prospect and retrospect means the time, the timing of when you collect the data, yes.

But to answer your question so observational study indeed can have both retrospective study and prospective study. They are just two different designs.

Q. Okay. So you see the prospective observational study is generally considered superior to retrospective observational study, correct?

A. Correct.

Q. Okay. And then on page 10, you also say that the design of the lowest quality is a retrospective observational study?

A. Yes.

Q. Do you see that?

A. Did I say --

Q. It's kind of at the bottom of page 10.

A. Yes. Okay. So to be clear, here I say the lowest quality that is of course I have a scope. I meant among, you know, if you have like three categories, you can, of course, further define them,

refine them into more categories. But I'm saying that randomized study, prospective observational study and retrospective observational study, this is -- among the three, this is of the lowest quality.

Q. Okay. And does that mean that retrospective observational studies do not have value?

MR. RODRIGUEZ: Objection, vague and outside the scope of the witness' opinions.

You can answer.

A. Yeah, well, I think you asked a similar question before saying whether this has value. Again, I don't provide opinion on that. They are studies. Those studies have, I call them low quality, because they have flaws in the designs, renders the conclusion to subject to be -- subject to all different biases. So they might not be reliable. But whether they have value or not, that's not what I'm -- I'm not providing opinion on that.

Q. When you say unreliable, you mean that they should not be relied upon?

A. I think that's an English word, unreliable has its obvious English meaning. It means that -- when I say unreliable, I mean that the conclusion is subject to a lot of biases.

So if you want to interpret that as

something strong or interpret that, interpret that those results -- you have to interpret the results with much caution and caveats. That's what I meant reliable, and they are -- those caveats or like those assumptions, if those are violated to any degree, then the result invalid.

So that's what I meant unreliable. They are just more subject to all sorts of challenges and, you know, biases.

Q. Okay. But you are not stating that they should not be relied upon?

A. I did not say that. But, of course, from common sense, if you have high quality, you want to make your decisions on high quality studies rather than low quality studies.

Q. Okay. So go to page 5 of your report.

A. Yes, I'm there.

Q. So in the second line under section 1, you say that: The main barrier to interpreting the association between the treatment and the outcome as a causal effect is the presence of factors that are associated with both the treatment and the outcome. These factors are commonly referred to as confounders or confounding variables or confounding factors.

A. Correct.

Q. Is the presence of confounders called confounding bias?

A. Yes. That's precisely why, again, as I later said, that's precisely the presence of confounding factors, that's precisely why retrospective study resolved before/after this data is viewed as low quality, because they cannot control at all the confounding factors.

Q. Okay. And that inability to control the confounding factors is inherent to a retrospective observational study?

A. It's inherent to -- so confounding bias is inherent to all observational study, whether it's prospective or retrospective. But prospective study do a better job in controlling those confounding bias by providing the before and after comparison.

Q. Okay. In going to page 6, you have a subsection called confounding bias. You state that randomized controlled trials eliminates all confounding bias.

Are there situations where that would not be true?

A. Again, that's -- we can get to too technical academic.

So, again, per se, if you have a study --

so raising the scope, if you do a randomized study that is because you flip a coin, so that is -- so for that specific study population, you're operating on the randomized study, yes, it does take into account, eliminate all the observed and unobserved confounding bias. So any confounding bias. So that's true. But then, of course, there are other things, if you want to put them a stretch, that a randomized study to a different population, then that's a different matter.

So, yes, so for the study per se, if you do a randomized study, operating on the population, the target population that you are operating the randomized study, yes, it eliminates all the confounding bias, and that's why it is regarded as gold standard.

Q. What if there's a very small sample size in a randomized control trial? Is it possible that would not eliminate all confounding factors?

A. Technically that's a different problem. So it still eliminate confounding bias just by design.

Small sample size is a different problem. Small sample size will give you larger extended error, or, in other words, the procedure, it will not be that precise. There is a statistical concept called standard error or variance. So that's a different one, like that's a second other problem.

So in confounding bias, the study design, randomized study design will admit no matter how small the sample size is. The small sample size, the key question is it will erode the procession of the study. So, again, statistically, it's like first order problem second order problem.

Q. Are there types of bias that can be present in randomized controlled trials?

A. Again, this is highly depend on the specifics, what do you mean? Like what type of bias?

As I mentioned, that if you want to stretch -- so the -- so confounding bias -- no, so the answer is so randomized study is subject to other type of bias, but not for the treatment effect for this population you study on. It's subject to other type of bias. For example, the randomized study, the population does not representing the general population, but that's a whole different matter.

And -- but always your first, again, like in FDA, if you want to have a new drug or treatment or medical device, you always first -- first order you do a randomized study. But that is understandable that it's not always feasible, but I won't go into that.

So your question, yeah, there are all sorts of different biases, but that is, again, we're talking

about first order problem and second order problem.

So why randomized study is prized is because the biggest problem to interpret the barrier from association to causation is confounding bias, and this randomized study is the most -- I mean, it's the single most effective design to admit that.

So that's a first order problem. Yeah, there are all sorts of other type of bias, but the second order problem and all the other observational study also subject too.

Q. Okay. So on page 6 and 7 of your report, you talk about different types of biases, and so selection bias, is that -- can that be present in a randomized controlled trial?

A. Oh, yes, that can be presented in randomized study or retrospect or prospective, but that's a different -- again, I say that's a second order problem, but yes.

Q. What about nonresponse bias?

A. Nonresponse bias, again, in randomized, yes, all of this study can be subject to that. But in randomized study, usually the way it's conducted, usually the nonresponse rate is controlled, because the study is controlled. So it's controlled by the -- highly controlled by investigated.

So it's usually subject to less of that kind of bias than the observational studies.

Q. Can recall bias be present in a randomized control trial?

A. Yeah, again, recall bias can present in all of these studies. But, again, because of the way that the randomized study is conducted, is well controlled by -- highly controlled by investigators, the occurrence of that is, the chance of that is much less than observational study, particularly for retrospective study. Because retrospective study often, like you don't design, because at the time you do the study, all these things already happened. So then that's a time lag.

That make it -- so in randomized study, you conduct the study. It's a prospective -- randomized study is prospective study. So you follow them. So, of course, there's much less chance of recall bias.

Q. Okay. What does it mean to mask or double mask a study?

A. Sorry, say it again, match?

Q. Mask or double mask, also known as blinding or double blinding?

A. Oh, I didn't say it here, yeah, double blinding, yeah, that is just a -- so that is you flip a

coin. Yeah, double blind, that means you don't -- so basically the treatment assignment, whether you get the true, the control or treatment is not known -- it's not revealed to the patient, and in some cases also not revealed to the people who conduct the study.

Q. So double masking means neither the participant or the conductors know who --

A. Correct.

Q. So if a randomized control study is not masked or double masked, can that induce bias?

A. Well, that can. I mean, I can give you a statistical lesson. Yes, it always -- like none of the study is perfect. Some are more imperfect. Others are less perfect. Yeah, so you're talking about double blindness. Yeah, the studies are not double blind. It's not always. So the problem of a double -- okay, so why do we want to do double blind study is try to reduce the chance of the so-called placebo effect. The placebo effect, like the psychological placebo effect, and, again, that is a possibility, like if you don't do double blind, there's a possibility that there's a placebo effect that will bias your result. But, again, comparing to other sorts of unmeasured confounding, there's a bigger problem. This is minor concern.

Q. On page 6, when you're talking about

confounding bias, you say: Therefore, in order to interpret the association between treatments and outcomes as causal effects in observational studies, one must assume that there's no unmeasured confounding factor. Such an assumption is untestable and is almost always untenable.

What do you mean by untenable here?

A. That means just almost always violated. There's always presence, in observational study, there almost always unmeasured confounded. And why I say it's untestable, because there's no unmeasured confounding. So it's unmeasured. How do you know whether there is or not? So, I mean, that's just common sense. That is like a standard in the literature.

Q. I'm not sure I followed. Can you explain it a little differently by what you mean? I get the untestable part, but can you tell me a little bit by what you mean untenable?

A. Untenable means, again, it's always violated to a certain degree, because you are basically saying -- so I'll give you an example.

If people say there's association between smoking and lung cancer, and then you calculate the association, so it's strongly correlated. But then if

you say smoking indeed caused lung cancer from observational study. If you want to make that statement, you basically say that well, then in my analysis, all the confounding factors, all the confounding factors that can affect both smoking and lung cancer has been collected and controlled in my analysis. And that is almost always weighted, because we can collect as much information about smoking and lung cancer, but there's always something missing, for example, like whether there might be genetic reasons, like your parents' -- like your parents' genes, or your parents' health, behavior, that kind of thing, you don't collect it.

So that's why I say it's always -- almost always they are -- you don't collect the whole universe of data.

So that's always -- this all matching confounding assumption in observational studies almost always violated. That's a consensus in the field, and that's actually the whole point like why people like me, a methodologist try to deal with this problem.

Q. Okay. So are you saying that observational studies can never be used to support treatment recommendations because there is a risk of confounding bias?

MR. RODRIGUEZ: Objection, asked and answered. Mischaracterization of testimony and report. And you can answer.

A. I think you asked this question or similar question many times. So I'll answer again.

I didn't make -- I didn't -- I don't come here or write my report to say that whether you can -- I didn't make the blanket statement to say that you cannot use observational study as the evidence for treatment recommendation. I purely said that they are different methods, different studies that can provide -- establish the treatment effectiveness of a treatment, and there are some better designs, some of worse design, and there's a reason I clearly describe here why there's a reason the confounding bias is the reason.

But, again, the answer has always been the same. I didn't say that because they are -- I didn't say that whether you should use it or not for your treatment recommendation, and that's not my scope. And I say again, I hope you don't ask this again, because my answer will always be the same.

Q. Is the concern that an outcome is the result of a confounding factor, rather than the treatment, mitigated by the number of studies

evaluating the treatment?

A. Can you say that again?

Q. So is the concern that an outcome is the result of a confounding factor, rather than the treatment being validated, is that concern mitigated by the number of studies evaluating the treatment?

A. I think your question mixed a lot of -- why I didn't first get it is I think you mixed a lot of statistical concepts.

So a confounding factor, a confounder is, by definition, is associated with both the treatment and the outcome, okay.

So I vaguely understand your question you're asking that whether this problem, this -- the problems of the existence of confounding bias or confounder is mitigated by the -- like the more study you do, like you are less concerned about that. Again, the answer of that is it depends on the quality of the study.

So if you give me one hundred studies, and one hundred studies are all very low quality, don't do a good job in controlling for confounders, if you give me 101, it doesn't matter, because they are all subject to the same problem. But if you give me a few high quality studies that did it good job in conjoining for

confounding bias then, yeah, it would be mitigated. But the sheer number, the number of study does not have nothing to do. The quality trumps quantity -- the quality trumps quantity here in terms of the studies.

Q. Okay. So you just gave the example it doesn't matter if it's a hundred or a thousand, so if you had a thousand observational studies with similar outcomes, you're saying that's still not as useful as having a few high quality studies?

A. Well, to answer your question, let's just look at the study I reviewed, right. Here I reviewed 80-some studies. So my -- based on my count, I think probably like 50 or, I don't know, 50 of them, I mean, not solvent. There's no solvent studies, but 50 of them are basically retrospective study resolved before/after comparison, and there are five prospective study that has the before/after comparison. And I explain why before after is important, because the before/after study provide you the most important confounder, which is the baseline measure of the outcome. So that's why it's regarded as better.

So I already said that, you know, you have 50, those 50 studies, even their own expert, even the own literature review say that they are low quality, and I'm calling for better prospective study, yeah, but

there are 50 of them providing the result. I would not, because they all have the same problem, they all subject to the confounding bias of, in this particular case, the confounding bias, particularly they're lacking the baseline. The most important confounder is that's the baseline outcome. They are basing that.

So they are all subject to that. Then, of course, you can do 50, you can do 100, you can do 1,000. That is just all subject to the same problem. Why do you repeat the mistake?

So that's why I don't -- I view that the evidence provide by a few high quality studies is better than 100 repeated, the studies low quality study repeat the same problem.

That being said, I do believe that it's better to have even -- the best would be to have both quality and quantity, means the best would be I have a high number of prospective studies. If you cannot do RCTs, that's fine. But if you can have a high number of high quality prospective study that all show consistent result, that's the best. But we don't have that, and actually the only ones they have have mixed results.

Q. When you say high quality studies, are you referring to randomized control trials?

A. No. I said that that's the best, but it's not always available. High quality, I meant that well designed, before/after retrospective study, if it's done nicely, done properly, yes, it can be viewed as high quality.

But, again, this is not the -- what is high quality or low quality is for any single study, of course, it's a subjective concept.

But here when I'm using the high quality and low quality I'm mostly talking about the, you know, from the perspective of whether it control for confounding bias. And why I say that is because confounding bias is the single most important barrier between the association and the causation, or that's the single most important barrier before you can establish the effectiveness of the treatment.

Q. So when you talk about high quality/low quality, you are not using those terms as they are used in the GRADE system?

A. Again, I'm not very familiar with the GRADE system. I don't know who decide that, and I don't know who decide that, and I'm not referring to GRADE system.

I already explained earlier in my case, I said there's at least three broad class of designs; one is randomized study, the other is prospective

observational study, and the last is retrospective observational study. And I say higher quality, low quality and lower quality is missing that. I have a hierarchy, one, two three, and I don't need to repeat that.

Q. And your hierarchy is yours and is not the GRADE system?

MR. RODRIGUEZ: Asked and answered.

You can answer.

A. Well, it's mine, but remember, I'm a national leading expert in inference, in study designs. So yes, it's mine. So I -- it's mine, but also I can tell you with confidence that is also the consensus in the field in terms --

Q. In terms of statistics?

A. In the field of causal inference comparative effectiveness research. That also include epidemiologist.

If you go out to any statistic, statistician, epidemiologist, ask them rank these three type of studies, they will give you exactly the same order as I just gave you. Randomized study is the top, and prospect is the second, and the retrospective third. So that's what I meant.

Q. I'm not trying to dispute that ranking. I

think I'm trying to understand, make sure that we're using the same language here.

And so my understanding from what you just said is you consider high quality studies to be randomized control trials or well designed prospective before and after observational studies, is that accurate?

A. Yes. Again, even there, yeah, randomized is still better than the other, yes, higher quality. But, again, even prospective before/after, you can still mess it up.

But, yeah, if it's like well, well studied, like it's well designed, yes, it can be considered as a high quality. But I didn't say every single prospective study is high quality, but it's still better than retrospective study.

Q. Are there ever reasons why a randomized control trial might not be ethical?

MR. RODRIGUEZ: Objection, outside the scope of this witness' opinions.

You can answer.

A. Yeah, I -- I don't think that whether a randomized study is ethical or not has anything to do with my opinions.

My opinion is about like whether the

treatment effectiveness is separate from the question of clinical recommendations?

A. Correct.

Q. Turning back to your report, how would you decide which WPATH assertions to evaluate?

A. How do I decide?

Q. Yes.

A. Oh, Orlando come to me and then he already, he gave me this numbers of assertions, asked me to examine that, and that's how I decide. I examined them one by one.

Q. Okay. Did you consider evaluating the assertions by the Endocrine Society and its clinical practice guidelines on the treatment of gender dysphoria?

A. Sorry. Can you ask that again?

Q. So the Endocrine Society also have clinical practice guidelines on treatment for gender dysphoria. Did you consider evaluating the assertions by the Endocrine Society in that guideline?

A. No, I didn't review. I'm not aware of that document, and I didn't review that.

Q. Okay. And in WPATH, in terms of care, are there recommendations made in that document?

A. I don't know. I don't remember. Again, I

focused on my opinion, or my expert report was focused on the specific assertions that was presented by Orlando to me, ask me to take a look.

So I didn't -- yeah, I glanced through that document, but I didn't remember everything that was said there or anything beyond that.

Q. Did you read the sections of the assertions that you were evaluating were in?

A. Oh, yes, I read, but, I mean, that's -- basically that's bullet points. Yes, I read those sections very quickly, but I think I actually did a good job in picking those assertions out is a good summary of the -- yeah, I mean that's paragraph.

Q. When you reviewed it, do you recall that there were recommendations numbered similar to what we just saw in the Pediatric Obesity Guideline?

A. I don't.

Q. Okay. And so you were not providing an opinion on the recommendations themselves?

A. No.

Q. Your chart at the end of your report lists Assertions 1 through 11. Are they numbered that way in the standards of care?

A. I don't know. I guess not. I mean, I don't know, because I didn't pay attention to the

standard of care.

Q. Okay. Were they numbered that way when Orlando gave them to you?

MR. RODRIGUEZ: Object, borderline on getting into communications between counsel and the retained expert.

A. I don't remember clearly. I believe that was the order he gave me, because I just then take it, and then it's like dealing with, paper -- I mean, it's dealing with one by one. So I didn't change the order or anything.

Q. Okay. In your report, you only include your evaluation of Assertions 1, 2, 6, 10 and 11. Why only those ones?

A. Say it again? So I only -- I believe that everything Orlando asked me to -- the assertion I indeed reply. Oh, 1 through 6? Why the -- okay. I need to look at the -- let me take a look.

Oh, did I say -- again, I don't remember clearly. I think the -- again, this I need -- actually, if I have a computer, I can see what is the original Assertion 7 to -- 7 to 9. I need to look. I don't know.

So if I didn't respond to that, it's -- oh, it's because --

THE WITNESS: Do you have the charge?

MR. RODRIGUEZ: Keep on looking through the rest of that.

THE WITNESS: Oh, okay.

A. Just give me some time. I look through the charge, then I know. Almost there.

So Assertion 7, so let me see, Assertion 7 says: Too often the agency's structure and personnel provide care are lacking in knowledge, training and capacity of care for gender diverse people.

So the paper -- well, oh, I see. So why I didn't, because this has nothing to do with my expertise, because my expertise was talking about -- it established the effectiveness or the safety of a medical intervention.

So for this assertion, I don't have opinion to provide, because I don't have expertise on this. So that's why I didn't provide.

So I think the same thing for the 8 and 9. So I provide the -- so there are a bunch of assertions, but I provide the opinions on the assertion that I feel that I have expertise on to judge.

Q. Okay. So you were not providing an opinion on Assertions 3 through 5 or 7 through 9?

A. I think -- yes. Yes.

Q. Okay. Well, let's go to Assertion 1 in the body of the report, not the chart.

A. Let me get there.

Q. I think that's page 11.

A. Yes, I found it.

Q. Okay.

A. Yes.

Q. Okay. So what do you mean when you conclude that studies failed to provide rigorous and statistical evidence on the benefits of quality of life and well-being of gender-affirming treatments? And apologies, I should have directed you to where that quote is. It's at the very end of your discussion of Assertion 1.

A. Okay.

Q. Page 15.

A. Yes. Yeah, so the end I said: I conclude, yeah, contrary to the statement in the assertion, these studies failed to provide rigorous and statistical evidence on the benefit of life and quality and well-being of gender-affirming treatments. Yes.

So your question? Can you repeat your question, please?

Q. Okay. Yeah. So I know earlier we talked about rigorous, and in this context, what do you mean

by rigorous?

A. Well, I explain that, okay. So if you look at the assertion, it says: There is strong evidence demonstrating the benefit, okay. So what do I mean by strong evidence? Okay. So I focused on that strong evidence. So as I go through my assertion, so I -- strong evidence, I show that they cited 21 studies, and so there are five prospective studies, which, again, in my GRADE, that is higher quality ones, and those have mixed results. And then they are now retrospective studies and others of lower qualities, and they have their lower qualities, and they are -- I describe all sorts of flaws in the design. And also I have, I said, there are also seven literature reviews.

So the literature reviews point out themselves that, you know, acknowledge the current available research based mostly on cross-sectional studies and call for -- so they acknowledge it's low quality studies.

So, I mean, rigorous, again, it's not rigorous. So they have this retrospective studies, and you would not cite those as rigorous, and the consistency part I already mentioned, because the higher quality prospective studies, actually the results show mixed. It's not always provide benefits

for that.

And 7, actually quite a sizable number of literature reviews point out, acknowledge the shortcomings and flaws in the current state of the research.

So that's what I meant it's not rigorous or consistent.

Q. Would there have needed to be a randomized control study here for you to find that there was rigorous and consistent statistical evidence to support this assertion?

A. If there's one that will add to the reader -- that will definitely add to the reader, there's none, which, again, that's not the end of the world, but they do have some higher quality ones, prospective ones. Unfortunately, they show the result of the benefit on quality of life is mixed. They find some positive, some are negative.

So, again, I don't -- I mean, I don't have an opinion. Like, again, it would be great if they have randomized study, but they don't have that, but I don't think that's the end of the world.

So I judge, I judge this, because they say that the assertions say there's a strong evidence. By my knowledge -- but by my examination of the literature

they reviewed, I find that's really a stretch, because the higher quality ones states actually mixed results.

I'm not going to repeat, but, yeah, that's what I meant.

Q. So if there are no randomized control studies, and the observational studies did not have mixed results, would you find that their assertion was supported by rigorous and consistent statistical evidence?

A. Can you say that again? Can you say -- sorry, I didn't hear clearly.

Q. Okay. So my last question was, you know, if there was a randomized control study, would you find that there was rigorous and consistent statistical evidence? And you said the lack of a randomized control study was not the end of the world. And so my next question is, you know, if there are no randomized control studies, and the prospective observational study has consistent results or not mixed results, would you find that there was rigorous and consistent statistical evidence to support this assertion?

MR. RODRIGUEZ: Object to speculation.

You can answer.

A. Yeah, this is speculation. So this is speculation. So I can only speak to the five, like the

studies they cited.

But to answer your question, yes, if they have like -- they have other prospective studies and well done, and they described nicely, it described very clearly the methodology and the result not mixed, I would feel a bit more confidence in -- like then, yeah, that definitely is some more strong evidence than what is currently presented here.

Q. And is there a set number of additional prospected numbers that would be needed?

A. No, I don't. Again, the better, the higher number, the better quality, it's better. But there's no number I can say well, you need five studies, you need ten studies. We don't have that.

Q. Okay. So I guess I'm trying to find where the line is for you, right. You've concluded that this is not rigorous and consistent. And so you must have some sense of what is, and like how do you make that determination? And it sounds like there's some gray areas. There's a lot of factors. How do you make that determination?

A. So the key thing, actually, I made my -- I made my opinion or my opinion, not like what you're saying that I don't need to make a line. I examine what is cited there, and what is cited there -- what

was cited there, the high quality ones already show you a mixed result. I don't think that I need to go further and say how much evidence I need to make this recommendation? Because I examine what is presented to me, what is cited by them. They cite it for a reason. They cite it for -- I assume they cite those papers to support the opinion. But what is cited there, if you close exam, provide mixed result that I think is enough. I don't need to -- for your question, it's beyond my scope to say oh, where should I draw that line? I judge -- I examine what is presented, what is cited by them, and thus result is not consistent, and I don't think that constitute a strong evidence for the -- strong evidence for treatment effectiveness, and that's what my opinion is about.

Q. And you know that this is not enough evidence to be considered rigorous and consistent, but you don't know how much would be needed to be rigorous and consistent?

A. Correct.

Q. Okay. So in here, you're talking about mixed results in the prospect area. Are you talking about is it the Lindqvist study?

A. That's one of the studies. I mention many. I also -- not many. I mean, there are just five,

right.

So Lindqvist is one. So, again, I went one by one. The other is before -- after.

Okay, so the other one, that is the Cardosa da Silva paper also find mixed results. Specifically they found that psychological side -- help and social relationship was significantly improved after SRS, but physical health and level of independence was significantly worse.

So that's another one. And another one is -- so another one, again, I don't know whether I need to read through, I mean, this says very clearly that the other one is a prospective study. So they focused on specific type of surgery, and they, the study does not provide no information about the fact of the surgery on general quality of life or well-being.

So then the other one, again, the five specific, the other one, the 2014 one, is focused on the safety and side effects of intervention, not quality of life. And the other one is prospective study, but that did not provide before/after comparison of the same patients.

So you can see that I list carefully there are five studies. There are five studies, and I mention that some of them do result mix. Some of them

simply do not provide information about quality of life, which is -- in the assertion, they say assertions about strong evidence of the quality of life or well-being.

So I'm saying that, okay, so there are two studies talk about that, the result is mixed. And then there are other studies that do not provide information for quality of life. So you cannot use that for evidence.

And the other one, the last one did not provide a before and after comparison of the same patient.

So that's why I clearly examined them one by one in five study, and then why I tell you that, you know, they're essentially among the five studies, there are two studies actually specifically talking about quality of life, and they are of high quality and the results are mixed.

Q. Okay. I'm trying to look at the document quickly. All right. So let's mark Exhibit 5, the Lindqvist study.

(The document referred to was marked Deposition Exhibit Number 5 for identification.)

A. Yes, I have it.

Q. And is this the study that you reviewed as part of your expert report?

A. Let me look at it. I believe so. It must be this, yes. Yes.

Q. Okay. And in the abstract, can you read the last sentence or the second to last sentence that starts with GRS?

A. The last sentence says: GRS -- so no, the abstract, they have background, they have methods, they have results, which one are you talking about?

Q. On the very first page where it says abstract, the last sentence of that section on the upper right side it says GRS.

A. Oh, GRS.

Q. Gender --

A. Yes, I read that: GRS lead to an improvement in general well-being as a trend, but over the long term, quality of life decreased slightly in line with that of comparison group.

Q. Does GRS refer to gender reassignment surgery?

A. Yes.

Q. Okay. And so this study, the people who conducted this study, did they conclude that gender reassignment surgery leads to an improvement in general

well-being?

A. Well, it said it proved as a trend but, but the key thing is but over long term, quality of life decreased slightly.

Q. Right. And what does it mean to decrease in line with that of the comparison group?

A. Well, he said that the comparison group also decreased, and the quality of life, yeah, it's -- I need to read more carefully what the comparison group, so who they compared to. I mean, it's one of the many papers. But the comparison group, I don't know what the comparison group is, without looking carefully. Is the comparison group -- I mean, they have a comparison group. I don't know that the comparison group is -- actually, I don't know here.

Oh, okay. So -- okay. So I'm now on page 225, and so the last sentence: Our findings on lower quality of life in transgender woman compared to the general population of women is in line with the same previous studies, and in contrast to others. However, one of those were performed on transgender men only.

So, again, they have a comparison. So I guess it's a comparison compared to the general population of women? I assume. Like that's my understanding, just quickly reading through this.

Q. Okay. And so you've described this study as having mixed results, why?

A. Well, because the first -- again, by mixed, I mean the initial, there's the increase, and then what I describe here then over the years at three years and five years, it's decreased. So that's mixed. It's not like by mixed I mean the trend of direction, the direction of the outcome is not one-sided, like maintain the same over years. That's what I meant, mixed.

Q. Okay. And you conclude that even though the comparison group had a similar decline with the same timing?

A. Again, that's their conclusion. They say it's similar. But if you look at the numbers, I don't know how similar that is, whether that's significant.

What they say is yes, that is in line, basically the comparison group, like, it's the general population, I assume, yes.

Q. So is it possible to interpret this as not mixed, because everyone is going to have a slight decline in quality of life at that time?

A. No, this is -- no, this is -- I cannot interpret that way, because this is about -- because the treatment in fact is on the population, on the

patient who received the treatment.

So you're talking about for this person, for those patients who receive the treatment, so then it's going up and then going down. You cannot stretch that to well, then there's -- like the fact is that there's indeed decrease. And what happened about the general population, I don't know, and that's not because they don't provide data for that. They are just saying oh, that is acknowledged by other, other studies.

Q. Okay. So do you disagree with the conclusion of this paper that gender reassignment surgery needs improvement in general well-being as a trend, but over the long term, the quality of life decreases slightly in line with that of the comparison group?

A. Not the conclusion. I take it at face value. I don't disagree. I mean, that's what they said, and then I just use that as one of the many papers. I mean, I don't know whether I agree. I take that at face value, this is what they said.

Q. But do you agree with it?

MR. RODRIGUEZ: Asked and answered.

You can answer.

A. Well, according to their study, according

to what they described, I agree with the conclusion. Again, they clearly try to spin it, make it sound more positive than it is. But that's what we all do when you write papers, because you get published that way.

But, yes, I mean, I think what their conclusion is supported by their data.

Q. You said they clearly tried to spin it. Where is that clear?

A. Where is that clear? So let's look at the numbers. Again, this is getting too details, but I can try.

So there's a decrease of general health. I'm looking at page 225. Again, this is one of the many studies. So I'm now doing on cite examine the paper for you. Where I say the main clinical -- so okay. So this is what I -- let me see, this is year one, okay. So that's year one.

Let's look at table two. So you have year two, table two that's individuals in this study. Okay, all of that year 0, 1, 3, 5, okay, then let's look at the general -- I see. I see what -- okay. Because they didn't provide data actually about the general population, so if you look at this study, is table two is providing information about all individuals in the study. They are on this through, you know, the

treatment. They didn't provide any information about the general population. There's nothing qualitative about their decrease. They purely just say oh, there are other papers. There's general decrease of the quality of life, and I don't, at least as far as I can see here, they don't provide information, quality of information, don't provide data to talk about the comparison group. So that's why I say that it's really a stretch, because they just say oh, there are other studies, but this paper did not provide that data, and they just say that it's inline, but there's no data supporting their argument, so, but, yeah.

Q. Do they --

A. Sorry, keep going.

Q. Go ahead.

A. No, I was just saying, it was very clear that -- but the trend talking about the patients in this particular study, as far as increase then decrease, that is clear. That is supported by the data by table 2. But then saying that is in line with the other general populations, that aspect, I don't see data supporting them. So that's why I see clear.

Q. Did you read this paper in your entirety in preparation for your expert report?

A. Yes, I read through that question quickly.

I focused more on the design and the data points. Yes, I read through that.

Q. Did you read fully all of the studies cited in your expert report?

A. Again, how do I define fully, right? I mean, I skim through them, and I focus -- I know the abstract, the conclusion. I skim through the -- very quickly the paper, but I focus mostly on the statistical methodology and the results and the numbers, that aspect. But, again, I guess that is entirety, but there's no -- I didn't check all the references and stuff, because that's -- that will take an infinite of time.

Q. So going back to WPATH Assertion 1 on page 11 of your report.

A. Okay. Yes, I'm there.

Q. Okay. So do you disagree that there are benefits in quality of life and well-being of gender-affirming treatment, including endocrine and surgical procedures?

MR. RODRIGUEZ: Objection, scope.

You can answer.

A. So, I mean, again, my opinion is about whether the references cited provide strong evidence for, as they claimed, for the benefit of quality of

life, and that's all my opinion is about.

So sorry, what are you asking?

Q. I was asking if you disagree that there are benefits in quality of life and well-being of gender-affirming treatment, including endocrine and surgical procedures, properly indicated and performed, and outlined by the standards of care?

A. As I said, they are papers that they cite that shows that -- again, those papers show, yeah, there is benefit in quality of life, but unfortunately, those studies are of low quality, and they are subject to all sorts of biases and the design flaws. So then how reliable that evidence is, that's questionable.

And so that's why I disagree with these assertions that say there's strong evidence, because this is at best pretty flimsy evidence.

Q. Okay. So the scope of your opinion is just as to whether or not there is strong evidence?

A. Yes.

Q. Okay. And you are not providing an expert opinion on whether there are benefits in quality of life and well-being of gender-affirming treatment?

A. No.

Q. Okay. Going to page 15, still in Assertion 1.

A. Yes, I'm there.

Q. On the fourth line down, oh, sorry, a little bit further down, I think roughly the sixth line, you said: None of the studies compare sexual assignment surgery with alternative treatment.

So what alternative treatment are you thinking of?

A. Oh, any. Like, so -- there so, again, this is also kind of academic in the sense that you compare.

So I say this it's because -- so they might be, again, gender dysphoria, I'm not an expert on that, but I'm saying that, for example, heart disease. There might be assorted different medications, right. So often we study -- in comparative effectiveness research, often we kind of do studies or trials, observational study, to see that among these different alternative, like four different medications, which one is the best? So that's what I meant by alternative treatments.

So here I say that, you know, this is -- they're not talking about, they're just talking about one particular type of treatment. And then I don't know whether there are other alternative treatments, but I'm just saying that in the statement, that here, they don't talk about -- like maybe their hormone

therapy. I don't know. But they don't compare like the surgery versus hormone therapy. Because, again, I over here, because later, in the later assertion there, the reason I'm thinking Dr. Ettner's point she made that, she made the one statement something like this is the only effective treatment.

So I think that is -- when I wrote that part, I was having that in mind. It's like well, all of these studies, particularly about having this treatment versus, I guess, not having this treatment, but not versus alternative treatment, for example, hormone therapy, things like that. So that's what I meant.

Q. But you're not aware of any alternative treatments?

MR. RODRIGUEZ: Object to the scope. You can answer.

A. Again, this is -- this depends on like, yeah, I'm not an expert of that. I don't know. I cannot say, have a blanket statement say that there's no alternative treatments or I am aware of that, because it obviously depends, there's different type of surgeries out there. I know that there's surgery versus hormone therapy. That's probably more or less -- and I also know there's different types of

surgeries.

So that's what I have in mind when I wrote that alternative treatment. I'm just saying that they are all -- all the studies here is comparing having this versus not, but not different treatment. You know, that's what I meant.

Q. So let's move on to WPATH Assertion 2 and looking at the bottom of page 16.

A. Yes.

Q. So you summarize that because the studies are all of low quality and they're subject to selection bias, nonresponse bias and recall bias, that they fail to provide rigorous and statistical evidence for the assertion, is that correct?

MR. RODRIGUEZ: Objection, incomplete -- incomplete recitation of the opinion. You can answer.

A. Yes, correct. I mean, you were basically reading out the rigorous and statistical sentence of mine. I reported on that point, yes. Yes, that's my opinion.

Q. Okay. And earlier we talked about nonresponse bias and recall bias, both being present in randomized control trials as well as observational studies, correct?

A. Correct.

Q. Okay. And so is it your opinion here that because there are no high quality randomized control trials, there could not be rigorous statistical evidence to support the assertions?

A. That's not my opinion. My opinion, as I said, that there are different -- there are different classes of study. Randomized control study is the best. If it's not available and not always available, then you resort to observational study. But even in observational study, there are good quality and poorer quality ones, higher quality or lower quality ones. And that I didn't say, I didn't say like there's no randomized study then -- like I didn't make my statements or opinion based on, entirely based on the lack of randomized study. Yeah, it is one of the weakness, but that is not the end of the world. I am more focused on actually the higher quality ones that are there. You know, there are very few high quality ones, and the results are mixed, and most of them are low quality ones. And I explain that, why they are low quality, what kind of bias they are subject to.

Q. And is it your assertion that the studies are subject to selection bias based on the fact that they are not randomized control trials?

A. No, that's not -- no, it's a different, it's a different aspect. So randomized, again, as I explain earlier, so why do you want to randomize study? All the classification of the quality of the study including inference.

The big -- like there are all sorts of considerations, but the biggest hurdle is confounding.

So randomized trial is good in dealing with that, okay. And so that was my point. And then as I said, the other type of bias, selection bias or nonresponse bias, yeah, those are all the studies subject to, but they are more of the second order issue. And also randomized subject is subject to those, but because of the ways randomized study is highly controlled by the investigators, the chance of those biases are much smaller than the observational studies.

So that's why randomized studies always prized, but I repeatedly said that without that, it's not the end of the world. There's still plenty of good well-designed observational study that, you know, can be valuable.

Q. You say that most of these studies are subject to selection bias. Why do you say that?

A. Well, because that's just the way it is.

Okay. I can go into details.

So one thing I can think of is many of the studies are from like a single, like a patient -- like a doctor from all the patients he has treated, or like they are locally in one hospital, one medical center over the years have done.

So again, in that case, then the patient, the patient population they consider might not be representative of the general population.

So that's what I meant. I mean, that's just a general, like if you read through the studies and I found that most of them have that kind of issue.

But I do note that they are studies, like -- I don't know whether it's here -- somewhere there's a study of like the Danish study of all the national registry over 20 years of the whole population. Yeah, that one I think, as I said, the quality is higher, because it's a more general population. But even that, it's still like focused on the Danish population. Whether that is representing the whole universe of transgender people, I don't know, and most likely it's not.

Q. Okay. And is your assertion that most of these studies are subject to recall bias based on the fact that some of the studies are retrospective?

A. Not some of them, yes. If they're retrospective, they are more likely subject to recall bias.

Q. Okay. So retrospective studies inherently have a degree of recall bias?

A. Yes, because, again, because this happened. You usually collect the data once things already happened. So that is -- you don't like collect the data -- the happening of the thing. Yes, it's inherent.

Q. And why do you conclude that these studies are subject to nonresponse bias?

A. Oh, well, nonresponse bias, as I mentioned, that when you have -- there's some patients who just do not provide data. And it's pretty well-known methodology that, you know, the people who do not respond to -- there's a reason they don't respond. They don't provide information.

So most likely they are different from the people who respond, who provide information.

So, again, or if you want to make them -- if you -- like only if you make the assumption that these two people, the people who respond versus people who don't respond are exactly the same, like if you want to say there's no nonresponse bias, basically you

are saying that the people who respond versus people who don't respond, they are the same. But that assumption most likely is, you know, not plausible.

So that is why I say that they're all subject to nonresponse bias. And I read in many of the studies, actually the nonresponse rate is pretty high, like they send out the surveys to the patients, and often the response rate is like 30 percent. That's really low.

Q. Okay. And are all types of study designs subject to nonresponse bias?

A. Yes, that's correct. But as I said, that some of the studies, like a randomized study, because you are -- you closely monitor them. Like you sign agreement, I think. So there are ways of improving that, but among all of the -- like the retrospective study, again, has the highest -- like is most vulnerable to that, because you have very little control over the response.

Q. Looking at the assertion itself, do you disagree with the statement that gender-affirming interventions are not considered experimental, cosmetic or for the mere inconvenience of a patient?

MR. RODRIGUEZ: Object to incomplete statement of the actual text.

You can answer.

A. This is not -- again, my opinion is -- so if you see that in what I wrote, I focused on the last sentence: They are safe and effective in reducing the gender, the gender dysphoria.

So my statement was not about like whether they are considered experimental, cosmetic for the convenience of patients, because, again, I'm not an expert in the medical practice of transgender people. So I cannot make -- I don't think I have the expertise to make a statement on that. But my statement is more about the second sentence, is they are safe and effective in reducing gender dysphoria, because that is -- fall into my expertise in comparative effectiveness research.

Q. And you are not providing an expert opinion on whether gender-affirming interventions are safe and effective at reducing gender incongruence and gender dysphoria, correct?

A. Well, I'm actually providing an opinion on whether this statement is supported by the references they cited, because they make this statement obviously based on, based on those references. And I want to examine that. And I find that statement is I disagree with that statement. Well, it's not I disagree with

the statement. My opinion is about that statement is not supported by the reference they cited.

Q. Okay. And so you are not providing an opinion on the safety and efficacy of gender-affirming interventions?

A. I don't, but if you read the sentence, they say they are safe and effective. And I say that this is not supported by the reference they cited.

So if you -- if you just not even stretch, if you continue on that argument, I don't think there's enough evidence to support that statement.

But, again, yeah, I'm not saying that whether this is -- whether those are safe and effective. I can only say those papers they cite do not provide evidence for this statement.

Myself don't have opinion on whether they are safe and effective, because I don't have enough evidence. At least all the evidence they provided, I would not make this kind of statement based on the references they cited. That's -- yeah. That is my opinion.

So based on the documents or the papers they cited, I would not make this kind of statement.

MS. NOWLIN-SOHL: Okay. We've been going for about an hour. Are you good to

keep going? Would you like a break?

THE WITNESS: Maybe five-minute break. Again, I drink too much water.

MS. NOWLIN-SOHL: Five-minute break. Sounds good.

(Off the record at 2:11 p.m.)

(On the record at 2:16 p.m.)

BY MS. NOWLIN-SOHL:

Q. We're back on the record. I'd like to move on with the WPATH number 10, which is on page 17.

A. Yes, I am there.

Q. Okay. So on page 18 in the middle paragraph at the end, you see the last point: Most studies do not discuss quality of life outcomes?

What is the significance of the fact that most studies do not discuss quality of life outcomes?

A. Okay. Let me look at the assertion first, just to refresh my mind.

So it says: Although different assessments result from -- okay. So the assertion says: Although different assessment measurements were used, the results from all studies consistently reported both a high level of patient satisfaction as well as satisfaction with sexual function. Although different measures -- assessment measurements were used, the

results from all studies consistently reported both high level -- oh, did I repeat? I think this is --

MR. RODRIGUEZ: Yes.

A. It's a typo.

So this was especially evident when used more recent surgical gender techniques. Gender-affirming, this surgery was also associated with low level -- low rate of complication and a low incidence of regret.

So then I wrote -- so you mentioned that I say most of the study is not -- which one -- did not provide -- which sentence -- did not provide quality of life. Did I -- sorry, I was just trying to -- can you repoint out like which sentence you are referring to?

Q. Yes, it's on page 18, the bottom of the middle paragraph.

A. Oh, bottom of the middle paragraph. Okay. Okay. Oh, most of the -- most studies use self-reported outcomes, instead of standardized instrument. And lastly, most study do not discuss quality of life.

Oh, so if you ask about the last sentence, most of the studies do not discuss quality of life specific to this assertion. They -- so, yeah, they talk about patient satisfaction and the satisfaction

with sexual functions.

So in that regard, yeah, they don't discuss quality of life. It's not a problem for this assertion. But also -- but I'm still reviewing the whole thing in the context of also the general forming my opinions. My general, like my overall opinion then in that case those studies that cited here don't talk about quality of life, then that is a concern, because my opinion was it was a part of that. It was about quality of life. Specific to this, to this assertion, this is not a concern.

Q. Okay. And then the bottom of page 18, the last full sentence you say: This body of literature supports the high self-reported satisfaction rate among the patients who underwent gender affirming vaginoplasty.

And how do you conclude that it supports that?

A. Oh, so are you referring to my sentence that says: But does not provide any evidence for the necessity or advantage of GAV comparing to alternative treatments? Are you asking me how I make that conclusion?

Q. It's that sentence, but I'm just asking about the first half of it right now.

A. Okay. So the body supports high self-reported among the patients who underwent -- yes, that is a fact, yes. That is a fact, yeah. This body of literature indeed support that, because that's a fact. And -- but the emphasis here is the self-reported, as I earlier discuss, self-reported also acknowledged by many of the expert in this field, in these papers.

So the self-reported outcomes often have -- it's also subject to some methodology flaw.

So in terms of this sentence, yes, that's just a state of fact, that's correct.

Q. You said that's a fact?

A. I mean, that's a fact that taking their numbers and their numbers says that consistent report, both a high level -- I'm saying the assertion, if you go back to the assertion, page 17, it said that results from all the studies consistently reported both a high level of patient satisfaction, so 78 percent to 100 percent, yeah, I mean, if you take that at face value, take that, then that is indeed a high satisfaction rate. I mean, that is a fact. A fact in the sense that the number stated there, and I trust that they didn't make up that number.

Q. Okay. And so you find that that body of

literature can support the high satisfaction rate even though there's no randomized control trial?

MR. RODRIGUEZ: Object to the mischaracterization of the actual text.

You can answer.

A. Yeah, here the English, I probably shouldn't say support. The body of literature basically just reports. I would say the body of literature report a high self-reported satisfaction.

So you can cross this supports, but I would use it as a reports, and that's what I meant the literature indeed reports that number.

Q. All right. And then the second half of your sentence says: But does not provide any evidence for the necessity or advantage of gender-affirming vaginoplasty comparing to alternative treatments.

Can you explain to me what you mean by that conclusion?

A. Yes. So as I earlier discussed to you, that -- so all of these studies is talking about this one particular surgery, GAV. So the study subjects are all the people who underwent GAV, and then they go to, as far as I remembered, the researchers surveyed those patients and asked whether they are satisfied, and, you know, their satisfaction and things like that. But it

didn't compare to people who could take alternative treatment, for example, hormone therapy or, you know, other -- I don't know. I only know hormone therapy.

So it did not compare GAV with anyone who takes say hormone therapy. Also, it didn't compare people who take GAV versus people who didn't take GAV.

So that's what I meant.

So you are just saying the statement, the study decided is focused on the people who indeed take this surgery, right, and then ask them whether you are satisfied. But there's nothing comparing there, because they didn't compare in the comparative world that had this person not take the treatment or had the person taken another treatment, like hormone therapy, what it would be.

So in that regard, that's like you can say oh, we've satisfied this, but there's no comparative statement. That's what I meant.

If there's no comparative statement, then you cannot say oh, this is the only one, that you must take this, because you don't know what will happen to those people that had they taken different treatment.

So the key thing is what I -- to summarize what I just said, it definitely means that these studies does not -- did not provide like any

information about comparisons, comparisons of this treatment versus control versus no treatment or other treatment, so that's what I meant.

Q. You mentioned hormone therapy as an alternative. Are you aware that the WPATH standards of care and the Endocrine Society recommend that people have at least 6 to 12 months of hormone therapy prior to undergoing gender-affirming surgery?

A. I'm not aware of that. Again, that's not my expertise. I say hormone therapy, it's just because when I review the literature, and I see people mention hormone therapy. I can -- I say hormone therapy, but when I put my sinus hat on, it would be treatment A versus treatment B. So hormone therapy would be treatment B.

Q. And do you know if anyone undergoes gender-affirming vaginoplasty or vulvoplasty without having hormone therapy?

A. I don't know. That sounds like a medical question that I cannot answer. I don't know.

Q. And your other conclusion does not provide any evidence for the necessity of gender-affirming vaginoplasty. Where does it say that that is what it is attempting to do?

A. Oh, it didn't. It's just, again, I just

rolled it. Probably when I rolled that, again, I was thinking -- well, okay.

So, yeah, it didn't, but I just make the statement because based on the fact -- based on the fact that none of this study comparing the treatment to alternative or to no treatment. So then once you don't have that, you know, you cannot say anything about this is necessary. Yes, this is not on the assertion. So this is -- I just roll that as if you don't have comparison, then you cannot, like there are a bunch of possibles, then you cannot say this is must, this is next.

I guess when I wrote that, I probably crossed -- was still thinking about some point of like this is the only effective thing or something like that.

So yes, you are right. In terms of in this -- specific to this assertion, they didn't, in this assertion, didn't say anywhere that it's necessary, okay. And I made -- yeah, I didn't -- well, this assertion actually is, yeah, I didn't make that statement. But, again, I was just based on the -- on the review I did, I add this -- I mean, I add this phrase there. That is accurate. The phrase is accurate. But, yeah, it is not, it's not stated in the

assertion, that is correct.

Q. Okay. And for the women who have been receiving hormone therapy and continue to have significant gender dysphoria, are you aware of any alternative treatments besides gender affirming surgery?

MR. RODRIGUEZ: Object to the scope.

You can answer.

A. No, that's, again, that's not my expertise. So I don't have answer to that. I don't have enough information about that.

Q. But you don't know if there are alternative treatments?

A. I don't know.

Q. Okay. Let's go to page 19, which has Ettner Assertion 1.

A. Yes, I'm there.

Q. Okay. And so in that -- your first paragraph after the assertion, about the fourth line down you say: As elaborated in my assessment of WPATH Assertion 1, the statistical methodology in the field of comparative effectiveness of SRS is not up to the long-established standard in comparative effectiveness research in medicine.

A. Correct.

Q. What is comparative effectiveness research?

A. Oh, comparative effective research, as I wrote in page 5 of my report, I said in medicine, the type of research used to evaluate the effects and safety of an intervention, it's broadly referred to as comparative effectiveness research. The statistical methodology for, again, here's a typo, for the comparative effectiveness research is generally -- belongs to the general statistics field of causal inference, which I am an expert on.

So basically in medicine and health studies, like the whole type of study I tried to establish the effectiveness, the safety, or efficacy of treatment is broadly referred to as comparative effectiveness research.

Again, you can see that there's emphasis, the two things: One is effectiveness, the other is comparative. So it's a comparison, yeah. That's what I meant.

Q. Okay. And did you mention there was a typo?

A. Yeah, the typo was -- so the last sentence says: The statistical methodology for quality of life, it shouldn't be quality of life. It's statistical methodology for comparative effectiveness research

belongs to the general statistical field called inference.

Q. Got it. That's the one we identified earlier?

A. Yes.

Q. Okay. So going back to page 19, what is the long-established standard that you're referring to in comparative effectiveness research in medicine?

A. Oh, well, as I said, that is -- that would be in the best case scenario, when available, you do a randomized experiment or maybe multiple randomized experiment. And when that is not available, you resort to observational study but well-designed observational study, and, for example, prospective before/after studies, and this like a more represent -- like a large study sample, things like that. And then like the other would be considered good or like acceptable. And then long-established, that, as I said, the consensus is then this kind of retrospect study resolved before/after measurement those -- and why you have also a lot of nonresponse as more sample size those of low quality.

So, again, that's what I meant by long-established standard and comparative effectiveness research in medicine.

Yeah, so like when you try to publish something, like about comparative effectiveness of treatment, then you expect to provide either randomized study or high quality observational study. So that's what I meant.

Q. And must that standard be met before a treatment can be provided to a patient?

MR. RODRIGUEZ: Objection to scope.

You can answer.

A. Not necessarily, because there's a lot of -- so, okay, different, different medical conditions like FDA. For example, if we are talking about new drugs and new medical devices, FDA would almost always -- well, most of the time would require a randomized study, right, but that's not the end of it, because they would do phase one, phase two, phase three randomized study, and then after approval, they will actually also do the post marketing, they call post marketing analysis.

Like, for example, we now all use Covid vaccine, right? We use Covid vaccine. So after Covid vaccine is proved, then actually there will be continuous study, actually to like in real world scenario. When Covid-19 vaccine is used, and then what's the population? So you go to continue. In that

case, you cannot do an online study, right. But you do an observational study for a large population, and then you calculate -- you study the effect in real life.

So, again, I think your question is like whether you should have kind of treatment provided before the -- sorry, can you rephrase? I cannot -- can you rephrase your question?

A. Yes. So you're talking about the methodology in the field of comparative effectiveness of sex reassignment surgery is not up to the long-established standard in medicine. And so I'm asking if that standard must be met before treatment can be provided?

MR. RODRIGUEZ: Objection, scope.

You can answer.

A. I think that's a different question from what I tried to -- from what my opinion is about, because I don't have -- I mean, I don't have answer to that.

What I can just say is the current state on the research on the sex reassignment surgery, at least based on the document they cited that I reviewed, do not -- because most of them are low quality, and the better quality ones have mixed results. Based on that, I would not -- I don't think that would meet the high

standards -- the high standard that people expect in, you know, he recommend -- I'm not going to recommend, just in reporting the effectiveness of those surgeries.

Q. So you are not providing an opinion on whether the evidence meets the standard in comparative effectiveness research, impacts whether treatment can be provided?

MR. RODRIGUEZ: Objection, scope.

You can answer.

A. Correct. I'm not providing opinion on that. I focused on Ettner's assertion. I focused on her assertion saying that studies show that gender-affirming surgery as safe and effective, safe and effective. And also she said that indeed for many people, this is the only effective treatment.

So my assertion is about whether she has enough evidence from those references she cited to support this two statement: One is whether they are safe and effective; the other is this is the only effective.

As I mentioned earlier, the only effective, that's kind of the necessity my understanding is. But when you say only effective, you have to compare, you have to can see that there's a possibility of alternatives. And that's one thing. And the first

sentence about safe and effective, I already said many times why, why the studies they decided do not meet the standard in providing rigorous and consistent evidence for that statement.

Q. Okay. And you're also not providing an opinion on what degree of statistical methodology is needed for a treatment to be included in a clinical practice guideline, correct?

MR. RODRIGUEZ: Asked and answered and scope. You can answer.

A. Correct. I'm not providing opinion on that. Again, I provided opinion on whether those assertions are supported or, yeah, by the documents they cited.

As I said, I would not, based on -- if I'm an expert in this field, which I am not in gender dysphoria, and if I'm an expert, and if I read this document -- no.

I'm a statistician, and I read those documents. I should say I'm a statistician. I read all the documents they cited, and then I would not make the statement as they make it. That's a better summary of what I want to say. Because I would say that there would not be enough evidence. I don't think there's evidence strong enough for me to make this kind of

statement.

Q. You disagree with Dr. Ettner's assertion describing the research as methodologically sound?

A. I disagree with that statement, and I think I said I don't think it's methodologically sound, and I already clarify why I think that many of the studies are fraud. And, again, that opinion has been -- that statement has been made in multiple, in large literature review in that field. So it's not -- yeah.

Q. Okay. And your first sort of comment after the sentence we were just talking about, comparative effectiveness research is that: There has not been a single randomized control trial. Is that the primary reason that you view the research is not methodologically sound?

A. No. As I said many times, that's not the end of the world, but it's one thing you can easily point out.

So my statement is -- my argument of -- basically my opinion consists of four things: First is there's no randomized study. Second is that's not the end of the world. You can still do good quality observational study, that is prospective before/after study. And then they don't -- and in the study they cited, they indeed have a few of them and the result

are mixed. And the third part of the argument is that in the vast majority of the study they cite are low quality retrospect study that subject to a lot of confounding bias and all sorts of biases. And the fourth component of my argument is that even in their -- many of their own large scale systematic review, literature review, they -- the expert are calling for better methodology or more prospective studies and call the current state of many of the studies of low quality.

So that's my -- like my statement has four components, and they are all integral.

So the lack of randomized study, that's the first component, but that's not the only component, and also that's not the only reason I made my statement.

So I think it's -- those four components are equally important.

Q. Are all cross-sectional retrospective studies methodologically unsound?

A. Well, all retrospect -- again, as a scientist, you don't make this kind of blanket statement. But what I can say, as I already repeatedly said, that you can have a prospective study if it's before/after data, that is far superior than retrospective study without before/after data. And the

reason of that is the confounding. The reason of that is the baseline measure of the outcome is often the most important predictive, the most important confounder out there. But retrospect studies do not control for that. So that's why it's low quality.

But I would not say that, you know, I would not blankedly(sic) say that every single retrospect study is garbage. No, that's not my point.

Q. Okay. But here, for this assertion, you're saying that the cross-sectional retrospective studies are not methodologically sound?

A. Yes, correct. In the studies they cite, I look at their methodology. They are, yeah, they're all lacking the baseline -- they're all lacking baseline measure of the outcomes, and they are subject to very severe confounding bias and they are not methodologically sound.

Yes, I stand by that statement in the context of what I reviewed.

Q. And are all retrospective studies subject to severe confounding bias?

A. Again, it depends on how much. So the answer is it depends on how many things you control for, how many confounders you control for, right. And as I said, that if in retrospective study, that if they

they use.

Q. Let's go back to the Ettner assertion, which is number 2 on page 20.

A. Yes, correct.

Q. In your report, you do not disagree with this assertion, correct?

A. Let me see. I need to read the conclusion. So I'm just reading it. The conclusion that is the gender-affirming surgery is the most appropriate treatment to elevate suffering of extreme gender dysphoria individuals still stands. 96 percent of patients who underwent that surgery were satisfied, and the regret was rare.

So what did I say? I said that: Immediately after the quote contained in this assertion, the authors themselves acknowledged. However, even today, this conclusion is based on methodologically less than perfect design studies. Okay. I guess I don't need to read it. But this is -- said very well what I think, right. And then this says, that this paper actually wrote that none of the study was a controlled one. So basically this paper described what are the -- why they say this methodologically less than perfectly -- less than perfect, right.

So the authors also acknowledge a few other methodological shortcomings, like attrition, or selection bias of the patient sample, which echo my critiques with respect to the WPATH Assertion 1. Yes.

Again, mine was based on -- so this particular assertion, so they analyzed that. So Ettner's assertion is -- yeah, so Ettner's assertion say that those -- they conclude that, and then I say that even as authors themselves that acknowledge the methodology is not perfect or is less than perfect.

Q. Okay. But you're not disagreeing with the statement?

A. I didn't provide opinion on this. Well, I didn't assess, because, again, as an expert, I'm a statistic expert, right. So then I can judge on the, again, from statistical aspect, whether this statement is based on like what kind of statistical strengths -- statistical evidence they have. And then I go to say -- well, I go to acknowledge, I go to find that the authors themselves acknowledge that those studies, you know, it's not -- the foundation is not most perfect, and I didn't direct -- in this assertion, I didn't directly say that -- yeah, I didn't provide opinion on whether that this is what I feel about this statement. Again, that's not what my opinion is about. My opinion

is about whether this is supported by the references, yeah.

Q. So looking at the Assertion 5 through 9, which is on page 21.

A. Correct.

Q. In your conclusion at the end of that section, does your use of the phrase rigorous and consistent scientific evidence, does that have the same meaning that we've discussed previously?

A. Correct.

Q. So if I were to tell you that 1 in 20 high schoolers in America die by suicide, would you say that America's high schoolers often die by suicide?

MR. RODRIGUEZ: Objection, speculation. You can answer.

A. That's how do we define often? Like my training is in mathematics, and you always say you need to give me a definition.

So if my definition is below -- is 10 percent -- if often is 10 percent, then it's not. If it's 5 percent, yes. So it depends on how you define often.

So my answer of this would be by definition tied to the definition of often, so I'm not going to speculate that.

Q. Okay. And how does one define often? How does one go about defining that?

A. Does that have anything to do with what we are talking about here?

Q. Yes.

A. So you're talking about a hypothetical about the high school, this like 1 out of 20. And as I said -- so as I said that if, I mean, I don't know, there's an English word, also if you want to go jargon, you can define this 5 percent or 10 percent. So --

Q. Does the context matter whether something is often or not?

A. Again, the definition matters. The definition matters, not the context. It's the definition.

Q. Okay. So if we say 10 percent is often, that would be often in every scenario that we can think of?

A. Again, you need to first define things. Like as a mathematician, we always first define a thing, especially if you want to attach, if you talk about numbers, you talk about often, yeah, the frequency. So the frequency, if I define -- if I say that by often I mean more than 10 percent, then that is often. But -- yeah, I really don't know how to further

elaborate this.

Q. Okay. How do you define often?

A. As I said -- well, as I said, that it's -- okay. So now I get to a point you say that depends on the context. I mean, I use this English word, and usually when I say often, it probably means something like over 50 percent.

So now I see your point. So, okay, so maybe it depend on the context.

So, again, this plain language or common sense often versus if you're talking about scientifically. So that's why in papers we don't write often, because this is like vague, non-defined terms, yeah.

Q. Okay. So let's move forward to page 23 and Assertion 11.

A. Yes.

Q. So midway through your paragraph, you comment on the Brown study, which reports that 5 percent of transgender inmates report that they had attempted or completed autocastration while incarcerated. And then you continue: This percentage supports the occurrence is relatively rare rather than often.

How did you define often and rare in that

assertion?

A. Well, okay, as I said, as common sense, I think often is probably 2 percent or 1 percent is not often, because that's 100 incidents. Then 100 times 2 or 3 times happen. I don't -- again, I don't -- I went from a layperson just a common sense. I would not say that's often.

Often, I would feel that is, if you say 10 percent or 20 percent, that sounds often.

So now I see your point. So you're going through this one, okay.

But as I said, I would not -- I mean, from a common sense perspective, regard 2 percent or 3 percent as often.

Q. Okay. Just a few minutes ago you said you would consider often as generally more than 50 percent?

A. No. As I said -- as I said, that is -- there's like if you want to go statistically, that I can say probably what you should write is probably Dr. Ettner or whatever should say by often, I mean this. And she didn't provide any of the numbers there. She just said often. I don't know what she mean by often. I'm just using the numbers I see at 2 percent or 3 percent, and I say that common sense, this is not often.

Q. But that assumes often is objective or consistent, doesn't it?

A. What do you mean by often is objective and consistent? Again, if you go to grab anyone on the street and ask them something happened 2 percent of the time, is it often? I bet you'd probably 99 percent of people tell you that's not often. That count as not often.

Q. Okay. We're talking about 5 percent here. So if I told you that 1 in 20 high schoolers, which is 5 percent, die by suicide in America, would you think that was often?

MR. RODRIGUEZ: Objection, scope.

You can answer.

A. Would I consider that often? So that's 5 percent, so that's another thing, okay. So high schoolers is a large population. So if you have say 10 million, let's say 1 million high schoolers, so 5 percent of that is, what, that would be 5,000, right? That would be 5,000. And then from that, yeah, that is -- so that quantity that is often.

Q. I think it would be 50,000.

A. 50,000 anyway, so that would be often in terms of the number, okay. But in rate wise, again, 5 percent, I don't know where to cut the line. But I

can tell you 2 percent or 3 percent is just not by any common sense would be viewed as common -- as often.

Q. Do you know what the rate of attempted or completed surgical treatment or autocastration is among non-transgender adults?

A. I do not, no.

Q. All right. Looking at Assertion 12, on the fourth line you say that: The WPATH study only provided qualitative studies, not any formal statistical meta-analysis.

Can qualitative evidence be useful in medicine?

MR. RODRIGUEZ: Objection, scope.

You can answer.

A. Of course it can. Of course they can be useful. But, again, this is -- what I wrote here is just a pure statement. It's -- of course qualitative would be better, because they provide -- they quantify things. I didn't say this -- their qualitative is not valuable. I'm not saying that.

Q. Okay. And then moving forward to the top of page 24, you are talking about other 9 studies. You say: The methodological problems include retrospective cross-sectional studies.

Is a retrospective cross-sectional study a

methodological problem?

A. Yes, I mean, again, as I said, retrospective cross-sectional means that there's no baseline measure of the outcome of the low quality, because it cannot control -- because it's subject to severe confounding bias, yes. So that is a problem. That is of low quality. And so the methodology did include that.

So it's one of the many weaknesses those studies are subject to.

Q. Do you disagree that researchers -- in looking at Ettner's assertion, do you disagree that researchers concluded that gender-affirming surgery positively affects well being, sexuality and quality of life in general?

A. As I said, that's not my opinion.

My opinion is that based on what they cited, the reference they cited, that those references they cited do not provide consistent or rigorous support for their statement.

In other words, if I'm a researcher, I look at those references. I would not make this statement, because I feel that the evidence is not strong enough.

Q. But you're not providing an opinion on whether the researchers conclude -- made that

conclusion?

A. Well, some of the researchers themselves said that they have called for -- again, they've called for this literature review. They called for better studies, better designs, and acknowledge the low quality of the current studies. So I stop there.

Q. So are you providing an opinion on whether the researchers concluded that gender-affirming surgery positively affects well-being, sexuality and quality of life in general?

MR. RODRIGUEZ: Asked and answered and scope. You can answer.

A. No, I didn't provide an opinion on that. I provided opinion on the quality of the study decided, and whether those studies that -- those studies would support the assertions they make.

Q. And so you are also not providing an opinion on whether gender affirming surgery does positively affect well-being, sexuality and quality of life in general?

MR. RODRIGUEZ: Scope, asked and answered. Go ahead.

A. Again, it's all of my scope, but as I repeated saying that, if I'm a researcher, and if I'm a researcher, and then I look at those studies, look at

those studies and decided I would not go to make that statement or those caveats, okay. Or I would acknowledge those assertions, or I would acknowledge the current state of the studies in this field do not provide consistent or rigorous conclusions and whether that equivalent to what you just said. That's not my -- that's a different question.

Q. Are you aware that the American Medical Association, the Endocrine Society and the American Psychological Association also put surgery in accordance with the WPATH standards of care as medically necessary treatments for individuals with severe gender dysphoria?

MR. RODRIGUEZ: Objection, scope, medical opinion. You can answer.

A. I'm not aware of that, because that's not my expertise.

Q. Okay. Do you think that they are wrong to do so, given your view that there is no rigorous and consistent statistical evidence on the benefits of sexual reassignment surgery?

MR. RODRIGUEZ: Objection, mischaracterization of testimony, scope and medical opinion. You can answer.

A. As I said, the treatment recommendations